ISSN: 2251-9246
EISSN: 2345-6221
pp. 163:169

Bidding Strategy on Demand Side Using Eligibility Traces Algorithm

Seyed Mohammad Ali Naseri Gavareshk*, Somayeh Hasanpour Darban, Amin Noori, Mahdi Besharatifar

Department of Electrical and Biomedical Engineering; Sadjad University of Technology, Mashhad, Iran
A.naseri1369@sadjad.ac.ir, S_hasanpour@sadjad.ac.ir, Amin.noori@sadjad.ac.ir, M.besharatifar143@sadjad.ac.ir

Abstract

Restructuring in the power industry is followed by splitting different parts and creating a competition between purchasing and selling sections. As a consequence, through an active participation in the energy market, the service provider companies and large consumers create a context for overcoming the problems resulted from lack of demand side participation in the market. The most prominent challenge for customers on demand side, is bidding strategy selection manner for attending in the competitive market. In this regard, they attempt to pay the least expense for purchasing the energy, while tolerating the least risk. In this paper, bidding strategy of service provider companies and the large consumers in the power market is proposed under the eligibility traces algorithm. In this algorithm, the demand side customers are considered as agents of Reinforcement Learning (RL). These agents learn through interaction with environment to bid such that earn the highest benefit.

Keywords: power market; bidding strategy; demand side; service provider companies; reinforcement learning; elgibility traces
Article history: Received 20-DEC-2017; Revised 29-DEC-2017; Accepted 15-JAN-2018.

1. Introduction

In the past two decades, in many countries, the electricity power industry under restructuring process is moving toward industrialization and competition. The main aim of restructuring is providing a condition in which power market determines the electricity price through increasing the competition and decreases the generation net expenses. This process, in electricity power industry is followed by separating different sectors and creating competition in generation and selling sections. Thus, selling the energy to the consumers is separated from distribution power networks operation, maintenance and extension.

In the competitive power market, small consumers do not tend to corporate in wholesale market or they are not allowed to intend in them. Therefore, there are companies which are responsible for providing energy to these consumers and distribution companies (Discos) and retailers are among them. These companies are often known as service provider companies. In order to satisfy their goals, these companies purchase the electricity from the wholesale market [1].

Discos are the owners of distribution power networks and responsible for the revenue. In restructured environment, the energy distribution companies purchase the electricity from the wholesale market in high voltage level. Then, the energy is sold to final consumers. It should be noted that in some restructured environments, retail companies are responsible for selling, instead of Discos [2]. The retail companies purchase the electricity from the wholesale market or purchase it through bilateral contracts with generation companies (Gencos) and sell it to the final consumers [3].

As a result, according to the intermediating or interfacial role of these companies, adopting an adequate bidding strategy is prominent and has a high influence on the company profit or loss.

On the other hand, large consumers often play an active role in the wholesale market through energy purchase. Under this condition, a competitive environment is created among service provider companies and large consumers [4].

At each time interval, if the amount of the real consumption on the demand side exceeds from that reported to the market, the independent system operator provides this shortage through auction in regulatory market, again. It should be noted that the independent system operator is responsible for preserving the balance between generation and consumption. This may have higher price in comparison with the competitive market. The service provider companies and large consumers cannot receive this additional amount from their customers, since this event was due to their incorrect performance in the market [5].

Through playing an active role in the market, service provider companies and large consumers create a context to overcome the problems resulted from lack of attending from demand side in power market [4].

In wholesale power markets, the service provider companies and large consumers are power market players who are looking for their profit maximization using proper strategies. Under such conditions, the most significant challenge for players is preserving their purchase price lower than their income and attaining reasonable profit.

Under restructured environment, the bidding strategy has become one of the most important issues for Gencos and consumers. The Gencos are competing to each other to sell the energy. According to the market principles, the competition winners are those who offer the least price. On the other side, the consumers are competing to each other for electricity purchase and those who offer the highest price, will win in the power market [6].

Consequently, an adequate and exact bidding will increase consumers' motivation to take part in the competitive market. The issues for selecting a suitable bidding strategy are:

On what basis, the consumers should offer to purchase?

In what range they should change their offer?

In what level the winner consumers should lower their price?

To what level loser consumers should increase their price?

All of these questions should be answered such that provide their demand power, while not losing. It should be stated that sometimes, the service provider companies may offer the price upper than the expected one to provide their own energy and not to encounter a problem in providing their subscribers. Sometimes, large consumers may risk reaching their goals, as well [1].

Selecting the right bidding strategy on the generation side is exclusively discussed. On the other hand, the buying offer strategy on the demand side is less studied. This issue is due to the fact that power markets on demand side, have not reached to the essential consistency for competing and the auction is usually performed unilaterally. Although, purchase strategy on demand side is an important part in the full power market and it is required to pay a specific attention and particular attempts.

In [4], a two-level mathematical model is represented for helping the large consumers in order to evaluate the proposed strategies to change the market clearing price to their advantage. The large consumers' desired bidding behavior in a special form of the power market is expressed in [5]. In [3], a method of investigating the desired strategy for retailers in short-time competitive markets is represented. The Genetic Algorithm (GA) is used to regulate the parameters for the best purchase strategy. A model based on the Cournot equilibrium is proposed in [7]. The Nash equilibrium approach for the midterm and long-term optimizations is proposed in [8]. Using this approach, no agents can unilaterally increase its profit in its strategy, The Monte-Carlo based algorithm to obtain the proposed curves for price estimation is stated in [9]. The consumers' bidding strategy using a risky method is mentioned in [6]. The information gap decision making theory in this paper, allows more severe decisions to be made against price frequencies.

One of the bidding methods, which is recently attained many attentions, is reinforcement learning methods. This algorithm learns based on the interaction with environment, through trial and error to select the optimal action to reach the goal.

In this paper, a new method of bidding strategy on demand side is proposed. In order to realize this issue, the eligibility traces techniques are used. In fact, the eligibility traces techniques are an interaction between Monte Carlo and Temporal Difference (TD), in which, having some numbers of Markovian Decision Process (MDP) can make the advantage of both algorithms.

In order to satisfy this aim, the paper is divided into the following sections: In section 2, customer's goal function for attending to the competitive power market is expressed. The eligibility traces method for bidding in the power market is stated in section 3. In section 4, the eligibility traces method implementation of bidding strategy on the demand side is stated. In section 5, bidding behavior study on the demand side in the competitive power market under a sample system is discussed and the conclusion is stated in section 6.

2. Market Structure

According to the figure 1, in the wholesale power market, the Gencos and the demand side have interaction with each other. In this case, the suppliers and the demand side offer their supply and purchase offers to the wholesale power market, respectively. After these prices are adapted by ISO, the power market is cleared.

The wholesale power market can be power pool or bilateral contract. In this paper, the market clearing structure is considered as power pool market and performed in a 24-hour period of time under Pay as bid legislation. Under Pay as bid pricing, the profit is allocated to the winners' market based on their offered prices.

The service provider companies and large consumers are players on the demand side, with the aim of maximizing their own profit in power market tenders. Consequently, the optimization cost function for demand side players is stated as (1).

$$F = \max\left[\sum_{i=1}^{H}(R_i) - \sum_{i=1}^{H} C_i P_i\right] \quad (1)$$

where, C_i indicates the costumers offered purchase price from the power network, P_i is the amount of purchased power, R_i, is the costumers' revenue function due to the purchased energy and H, indicates the number of the daily market implementation hours.

On one hand, we have no access to the revenue function and the manner of earning revenue by players. On the other hand, in this study, the aim is not the buyer's manner of earning revenue resulted from the won energy. So, the cost function is modified to the least purchase cost of energy as (2):

$$F = \min\left[\sum_{i=1}^{H} C_i P_i\right] \quad (2)$$

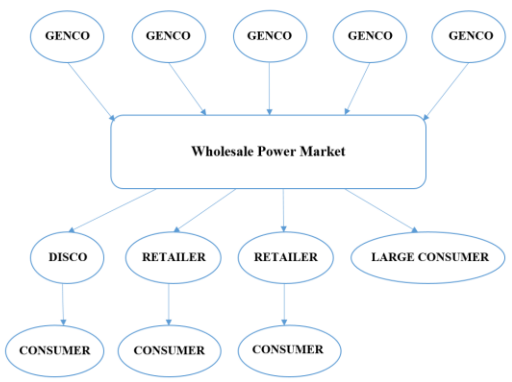

Fig. 1. The wholesale power market

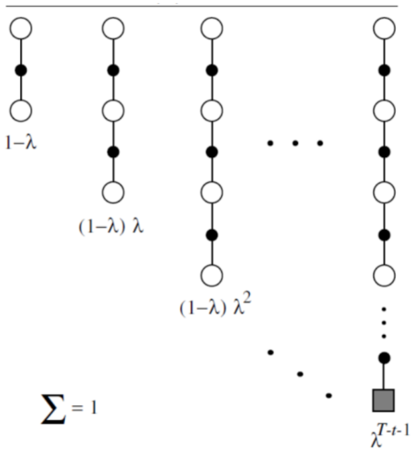

Fig. 2. Eligibility traces algorithm

3. Reinforcement Learning and Traces

RL is inspired from a psychological theory, resulted from analyzing animals and artificial systems behavior. In RL, the agent learns which action is better in each situation. In fact, the RL is a mapping from state space to actions space, so the reward signal is maximized. In RL, the agent is not told which action to select. The agent must find the action which results in more rewards, through its interaction with the environment, in the long term [10,11].

One of the RL solving problem is eligibility traces algorithm. The eligibility traces are combination of forward view of total rewards and backward view of errors. In forward view, like $TD(\lambda)$, averaging n-step backups is done. This average contains all the n-step backups each weighted to λ^{n-1} where $0 \le \lambda \le 1$ [10]. The resulting backup is toward the return, that is, the return of $TD(\lambda)$ is obtained as (3):

$$R_t^{\lambda} = (1-\lambda)\sum_{n=1}^{\infty}\lambda^{n-1}R_t^{(n)} \quad (3)$$

Figure 2, shows that a normalization factor of $1-\lambda$ and weighting factor of λ. If λ is assumed to be zero, then $TD(\lambda)$ is converted to the TD method and if λ is assumed to one, this method is converted to Monte-Carlo. Figure 2 demonstrates this weighting sequence. On each step, t, the update rule is as (4):

$$\Delta V_t(s_t) = \alpha[R_t^{\lambda} - V_t(s_t)] \quad (4)$$

The backward view of $TD(\lambda)$ provides an incremental mechanism for approximating the forward view. The backward view of

$TD(\lambda)$, there is an additional memory variable associated with each state. The eligibility traces for all states decay by $\gamma \times \lambda$ and the eligibility trace for the one state visited on the step is incremented by one as (5):

$$e_t(s)=\begin{cases}\gamma\lambda e_{t-1}(s) & \text{if } s\neq s_t\\ \gamma\lambda e_{t-1}(s)+1 & \text{if } s=s_t\end{cases} \tag{5}$$

where, γ is discounting factor. The trace in eligibility traces, indicates the degree to which that state is eligible for undergoing learning changes and should reinforcing event occur. The TD error for state value prediction is as (6):

$$\delta_t=r_{t+1}+\gamma V_t(s_{t+1})-V_t(s_t) \tag{6}$$

In the backward view of $TD(\lambda)$, the global TD error signal, triggers proportional updates to all recently visited states (7):

$$\Delta V_t(s)=\alpha\delta_t e_t(s) \tag{7}$$

This update (eq.7) could be done on each step or could be saved until the end of the episode.

4. Eligibility Trace implementations

In this paper, the intention is to investigate the bidding strategy behavior on the demand side, including service provider companies and large consumers. Each of the market players during the energy purchase bidding, attempts to maximize its own profit and its customers' satisfaction. For solving the bidding problem, it is needed to define the reinforcement learning elements in considered problem. Each of service provider companies or large consumers is indicating the agents that compete for achieving their own goals in the market environment. The states in each hour are determined based on market clearing price.

The state space is divided into the equal sections between lowest market price and the agent initial price. Each agent makes decision for offering its own price. The action space is divided into the equal sections between the lowest market price and the offered purchase price. At each time, the agent behavior is defined based on the optimal policy it selects. In this work, the optimal policy for action selection is based on the Simulated Annealing (SA) algorithm.

The aim of the algorithm is investigating the balance between exploration and exploitation during the eligibility traces algorithm. This algorithm specifies the policy of the agent to select the next action in order to get as much reward as possible. The action selection steps in SA are as figure 3. Also, the value of the reward for each agent is considered equal to the benefit of each agent.

The service provider companies and large consumers are looking for providing lower prices in the market in order to win in the market while paying less for energy purchase. As previously mentioned, this offered price may change to the price higher than the expected extent in order to reach the considered goals and in contrast, no profit is allocated to them.

Table.1.
The characteristics

Customer	*P_{Total} (MW)*	*Step number*	*MC ($/MWh)*	*Customer*	*P_{Total} (MW)*	*Step number*	*MC ($/MWh)*
			110				100
1	60	3	100	4	50	3	90
			95				80
			100				110
2	50	2	98	5	60	4	108
			95				95
			90				90
3	60	4	88	6	60	2	95
			85				90
				7	25	1	100

1. Initiate arbitrary all $Q(s,a)$ values.
2. Select an action $a_r\in A$ randomly
3. Select an action $a_p\in A$ based on policy:

$$a_p=\arg\max_a Q(s_{t+1},a)$$

4. Generate a random number $\in(0,1)$
5. Select and perform action $a_n\in A$ as follows:

$$a_n=\begin{cases}a_p & \text{if} \quad \xi\geq\exp\left[\dfrac{Q(s_t,a_r)-Q(s_t,a_p)}{Temperature}\right]\\ a_r & Otherwise\end{cases}$$

6. Recalculate Temperature by temperature dropping criterion.

Fig. 3. SA algorithm for policy updates [12]

After determining the reinforcement learning components, in order to obtain the proper initial values of the state-action table, the initial learning is performed for the number of considered learning period. After the initial learning period (offline learning) is completed, the initial values of Q-table for the main bidding (in an online mode) are used. At the end, the agents bidding algorithm in the market environment is executed according to the figure 4.

5. Numerical studies

In this section, the bidding strategy of 7 energy purchase customers in a sample system in day-ahead market is analyzed. The sample system characteristics are provided in Table 1. Since the generation costs are variants, it got possible to use step bidding method for customers (and suppliers) in various power markets all around the world.

ISSN: 2251-9246
EISSN: 2345-6221

1. The initial learning, initializing the Q-table and $e(s,a)=0 \quad for \quad all \quad s,a$
2. Repeat (for each episode):
   A. Initial offering and the initial state
   B. Repeat (for each step of episode):
      - Take action a, observe r, s' ($e.g.,$ $Simulated\ Annealing$)
      - Offerings collection by ISO and market clearing
      - Benefit calculation and receive
- $\delta = r + \gamma Q(s',a') - Q(s,a)$
- $e(s,a) \leftarrow e(s,a)+1$
- For all s,a
  $Q(s,a) \leftarrow Q(s,a) + \alpha\lambda e(s,a)$
  $e(s,a) \leftarrow \gamma\lambda e(s,a)$
- $s \leftarrow s'; \ a \leftarrow a'$
- until s is terminal

Fig. 4. The agent bidding procedure

In this method, the customers divide their energy capacity demand into different steps and sort the prices non-incrementally, to have more chance of winning in the market. In this case, the agent offers lower prices for the final steps that less energy is allocated to them, such that rejection or acceptance of the steps in the market, have no impact on its benefit. It should be mentioned that at each hour, the customers offered capacity at each step is variable. Actually, at each hour, the customer considers whether increase or decrease the step capacity in comparison with the same period of time on the previous day. In this manner, the customer enhances its opportunity to earn more benefit.

In figure 5, the amount of the energy provided by suppliers for participating in the market is represented.

In order to explore the bidding strategy, behaviors of customers 5 and 2 are considered. These customers offered their required capacity in various numbers of steps and different prices to the power market. Results are depicted in a 24-hour period.

Figures 6 and 7 show the energy market clearing price and the offered prices of both customers. The customers must offer the price higher than market clearing price. The higher is the offered price, the more chance the customers have to win in the market. It is not wondering that by this method of bidding, the customer should pay more for it. This behavior depends on the customer's policy. Sometimes, the customer gives up earning benefit for the sake of achieving its other aims. For instance, a service provider company is forced to adopt this policy in short period of time, in order to keep its customers. But over time, it can learn how to pay the minimum price to the networks for purchasing the energy.

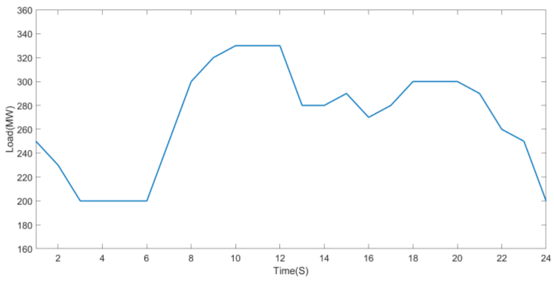

Fig. 5. The amount of the energy provided by suppliers for participating in the market is represented

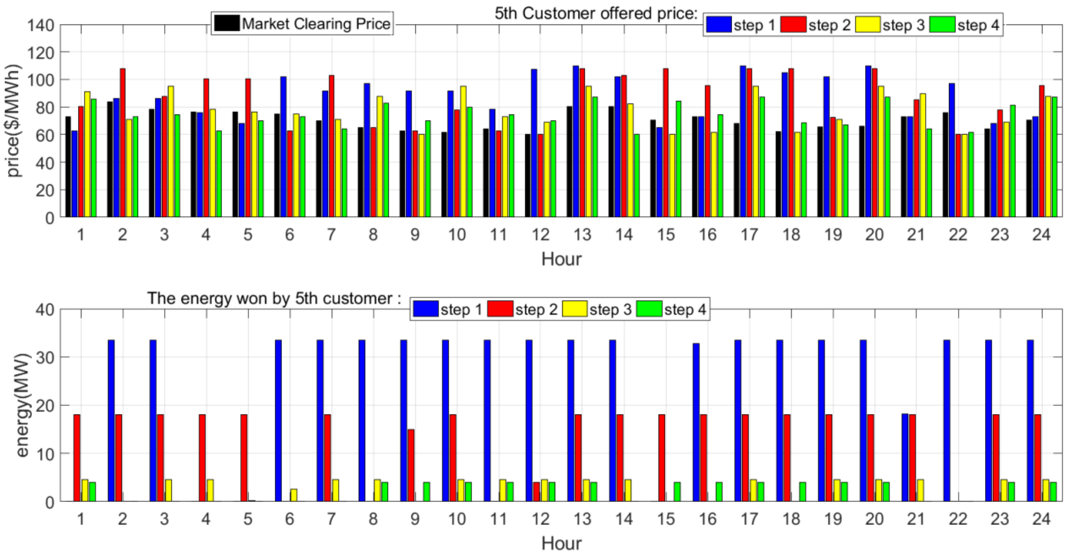

Fig. 6. The offered price and the purchased energy by the 5th customer

Fig. 7. The offered price and the purchased energy by the 2nd customer

According to figure 6, the customer number 5, can supply more share of its demand from the competitive market, through several biddings. At the first hours of the competition, the customer offers its final steps in low prices. Since, the selected capacity is small for this step, its winning or losing in the market has no impact on its income. In fact, this customer allocates the final steps to market situation identification. It is observed that after some hours, it can learn how to offer the final steps, such that it can make money and identify the market, as well.

In figure 7, the 2nd customer, participates with 2 steps in the market to win and supply more share of its demand energy. This customer, due to the lack of the variety in bidding, and risky participation in the market, the customer did not perform well at the first hours. Over the time, the customer became experienced and offered high prices in the market, which sometimes causes its costs to increase. In fact, as previously expressed, the customer gave up the benefit earning for the sake of achieving its other objectives.

In Table 2, the amount of demand energy, the purchased energy and the purchase cost by each of the customers 5 or 2 are represented for the first 24 hours in the competitive market. As can be observed, in comparison with 2nd customer, the 5th customer can supply more share of its demand energy from the competitive market, at the first 24 hours. The 2nd customer paid more money for this amount of purchase.

According to Table 3, over the time, the customer number 2 could have a more intelligent bidding through trial and error. In this case, it supplied more share of its demand energy from the competitive market, therefore paid less money. It is not wondering that the customer cannot supply its demand energy from the market. As a consequence, it is forced to purchase the demand energy from the momentary market and pays more money.

Table.2.
The market results at first 24 hours

customer	*the amount of demand energy (MW)*	*the purchased energy (MW)*	*he purchase cost ($)*
5	1440	1131.6	17062
2	1200	844.12	11738

Table.3.
The market results at second 24 hours

customer	*the amount of demand energy (MW)*	*the purchased energy (MW)*	*he purchase cost ($)*
5	1440	1184.4	14288
2	1200	1023.8	11860.3

6. Conclusion

In this paper, the eligibility traces algorithm was employed for the demand side bidding in the competitive electricity market to maximize the customers benefit, including distribution companies, retailers and large consumers, in the market tenders. The simulation results revealed that employing this algorithm, the customer's benefit was maximized. Also, it was corroborated that using this strategy, the customer could supply more share of its demand energy in the competitive market. Also, it was demonstrated that the customer could earn more profit by dividing its demand energy into several steps and offer them to the power market. In fact, the customer intelligence in selecting this strategy caused it to offer small share of its capacity to the market, in lower prices and resulted in the market clearing price reduction. This caused the customer to supply more share of its demand energy from the power market.

References

[1] Ya'an Liu, Xiaohong Guan, “Purchase allocation and demand bidding in electric power markets,” IEEE Transactions on Power Systems, vol. 18, 2003.

[1] Amir Safdarian, Mahmud Fotuhi-Firuzabad, Matti Lehtonen, “Integration of Price-Based Demand Response in DisCos' Short-Term Decision Model,” IEEE Transactions on Smart Grid, vol. 5, pp. 2235-2245, July 2014.

[2] R. Herranz, Antonio Munoz San Roque, José Villar, Fco Alberto Campos, “Optimal Demand-Side Bidding Strategies in Electricity Spot Markets,” IEEE Transactions on Power Systems, vol. 27, 2012.

[3] S. Jalal Kazempour, Antonio J. Conejo, Carlos Ruiz, “Strategic Bidding for a Large Consumer,” IEEE Transactions on Power Systems, vol. 30, pp. 848-856, July 2014.

[4] A. B. Philpott, E. Pettersen, “Optimizing Demand-Side Bids in Day-Ahead Electricity Markets,” IEEE Transactions on Power Systems, vol. 21, 2006.

[5] Kazem Zare, Antonio J.Conejo, Miguel Carrión, “Multi-market energy procurement for a large consumer using a risk-aversion procedure,” IEEE Transactions on Power Systems, vol. 80, pp. 63-70, January 2010.

[6] J. Contreras, O. Candiles, J. Ignacio de la Fuente, T. Gomez, “A cobweb bidding model for competitive electricity markets,” IEEE Transactions on Power Systems, 2002.

[7] C. A. Díaz, J. Villar, F. A. Campos, and J. Reneses, “Electricity market equilibrium based on conjectural variations,” Elect. Power Syst. Res., vol. 80, 2010.

[8] D. Menniti, F. Costanzo, N. Scordino, and N. Sorrentino, “Purchasebidding strategies of an energy coalition with demand-response capabilities,” IEEE Trans. Power Syst., vol. 24, no. 3, 2009.

[9] R. S. Sutton and A. G. Barto, Reinforcement Learning: An Introduction, Cambridge, MA, USA: MIT Press, 1998.

[10] Chunming Liu, Xin Xu, Dewen Hu, “Multiobjective Reinforcement Learning: A Comprehensive Overview,” IEEE Transactions on Systems, Man, and Cybernetics: Systems, vol. 48, 2014.

[11] Zhigang Liao, Ly-Fie Sugianto, "Using Q-learning to model bidding behaviour in electricity market simulation", IEEE Symposium on Computational Intelligence in Multicriteria Decision-Making, 2011.